Characterization of Extreme Hydroclimate Events in Earth System Models using ML/AI

Authors

Katrina E. Bennett, Satish Karra, Velimir V. Vesselinov
Earth and Environmental Sciences, Los Alamos National Laboratory

Focal Area(s)

(1) We put forward the concepts of data assimilation enabled by machine learning, AI, and advanced methods including experimental/network design/optimization and unsupervised learning applied to downscale information within Earth System Models (ESMs).
(2) We discuss predictive modeling through the use of AI techniques and other tools to design a prediction system comprising of a hierarchy of models (e.g., AI-driven model/component/parameterization selection) to improve the characterization of extreme hydroclimate events in ESMs. The AI-based models will run five-six order of magnitude times faster, yet will provide similar accuracy, allowing us to provide range bounds on uncertainty faster and thus enabling faster extreme event identification.
(3) Further, we consider the insight gleaned from complex data (observed/simulated) using AI, big data analytics, and other advanced methods, including explainable AI and physics- or knowledge-guided AI to improve the characterization of extreme hydroclimate events in ESMs.

Science Challenge

Hydroclimate extremes and perturbations are becoming increasingly common and intense at regional and global scales[1], and represent a major threat to national and global security. Future climate changes may allow for unprecedented rapid energy development in many regions[2], with reverberations to economy and society[3-5]. Concurrently, climate-change-induced increases in extreme events may hinder energy development, and critical energy infrastructure may experience severe damages as climate tipping points are exceeded[1]. Perturbations lead to cascading effects that ultimately impact regional stability. As these destabilizing impacts occur, the system moves into a ‘no analog’ future where the historic period may be a poor indicator of the magnitude, return interval, and timing of future events[6], from which recovery is difficult if not impossible. Despite this, some of the most vital tools to examine the climate change impacts, for instance, Earth System Models (ESM), do not replicate extreme events adequately. The recent advances in ML/AI show promise in filling this gap within ESMs. Our overarching vision is that ESMs require: 1) an improved research framework to integrate AI-assisted assimilation of observed extremes, 2) physics-informed machine learning for faster predictions of model/component/parameterizations, and 3) AI-informed data analytics to detect, characterize, and assess uncertainties in extreme hydroclimate events, addressing focus areas 1-3.

Rationale

Extreme hydroclimate events are typically estimated using downscaled climate model outputs, executed offline within land surface models to characterize and assess their impacts (e.g., drought, flooding) in the medium to long-term (50-100 years)[7]. Alternatively, we can use high-resolution, fine-physics models to forecast the impact of extreme hydroclimate events in the near-term future at fine spatial locations (e.g., forest fires). The fine-scale approach is time-consuming and relies upon tiers of pre-/post-

DOI: https://doi.org/10.2172/1769685
DOE BER, AI4ESP 2021 #AI4ESP1011 (https://www.ai4esp.org)

processing, data archival, collection of observations (where extremes are not represented or are limited in their numbers in the records), and the application of sophisticated statistical methods for identifying extreme events[8,9]. Rather than using tiers of models and processing/analysis steps, embedding some of these approaches directly within ESMs will revolutionize the efficiency and efficacy by which researchers understand extremes, and will vastly improve mitigation and adaptation approaches to manage extreme hydroclimate events better.

Although changes in extreme events are occurring regionally and globally, there is a gap in the ability of regional models and ESMs to represent extremes because these models fail to capture the physical mechanisms governing changing land cover and land use at scales to characterize responses accurately in space or time[1,10-12]. Therefore, alterations and feedbacks to the regional and global climate system from extreme events cannot be adequately captured and the impact of shifting frequency, intensity, and reoccurrence of extremes is currently not well known[13]. Additionally, observations of extreme events are rare by nature, and thus, the methodology to determine changes in events must be carefully considered to assess impacts correctly and present reliable results for decision-making purposes[14]. Finally, since the current science of extreme events includes local perturbations, it typically considers impacts on easily obtainable (e.g., temperature) but highly uncertain (e.g., precipitation) singular/independent variables, and disregards cumulative shifts in concurrently changing hydrological variables (e.g., forest fire followed by an atmospheric river event).

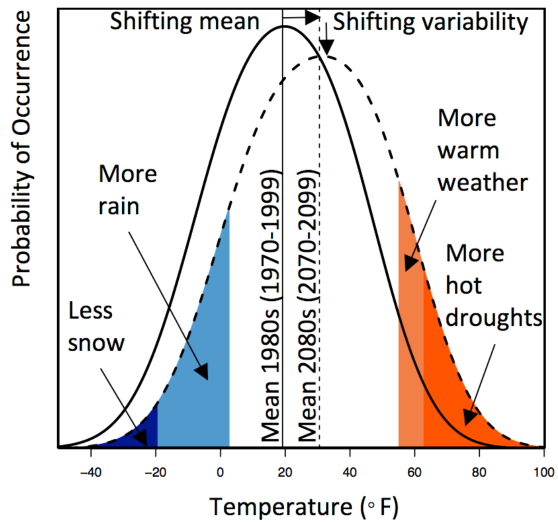

Figure 1. Shifting mean and variability in temperature and what this means for rain, snow, warm weather and hot droughts. Colorado River basin, USA.

Understanding singular extreme impacts is insufficient to define the problem or capture the range of responses and associated uncertainties that lead to the no-analog future and a decline in ecosystem function. The state-of-the-art science, therefore, completely misses the risk associated with concurrent extreme events that are a) interlinked, b) result in strong feedbacks to regional and global climate, and c) can exacerbate threshold behavior leading to new equilibria (Figure 1)[15,16]. Thus, a critical gap exists in predicting vulnerabilities and their impacts at regional and global scales in support of DOE's energy and environmental missions.

Current methods are hierarchical, independent, and require much manual work to carry them out (i.e., Coupled Model Intercomparison Project-level scale projects and databases, which are time-consuming and inefficient). A new paradigm in ESMs is to process observations using AI-assisted assimilation to define extremes, and compare/contrast their occurrence and character within observations to the ESM-derived events for in situ downscaling. Additionally, the use of reduced-order physics models embedded as sub-models in ESMs and coupled with ML reductionist analytical methods to assess extreme events and characterize their distributions will vastly improve parameterizations to better capture these events and represent them within ESMs.

Characterization of Extreme Hydroclimate Events in Earth System Models using ML/AI

Narrative

The development of ESMs capability to identify, characterize, parameterize extreme hydro-climate events will revolutionize the field of earth sciences by allowing for actionable science tools from ESMs towards mitigation and adaptation to these disruptive and destabilizing events. Observations of extremes exist in many different global databases (e.g., GHCN, SNOTEL, ARM)[17-19]. These observations can be automatically collected/assimilated using AI-assisted techniques, such as unsupervised ML approaches. Unsupervised ML approaches are known to identify patterns, key features, and dominant signals in complex data[20]. Dimensionality reduction methods such as principal component analysis (PCA) or more recent advances such as non-negative matrix/tensor factorization methods can be used to downscale and identify key signals and features of extreme events[20,21]. One of the most important components is to automate a definition at which the ESM will determine it requires the assimilation components to ensure that assimilated data does not result in other unintended consequences or physical changes within the model itself. Once these methods provide downscaled signals and features, ML scientists can work with climate scientists to interpret and define these outputs[22,23].

Another way that ML/AI can significantly improve existing ESM workflows is through reduced-order models that use ML. Instead of running computationally intensive fine-scale models, ML/AI models that mimic the underlying physics and are 5-6 orders of magnitude faster, can be used[24]. For instance, methods such as ensemble and multilayer perceptions have been shown to provide excellent (accurate yet fast) reduced-order models for complex nonlinear physics[25]. Reduced-order physics models embedded as sub-models within ESMs coupled with ML reductionist methods to assess extreme events and characterize their distributions, and for improved parameterization to better capture these events and represent them within ESMs.

ML/AI supervised techniques such as convolutional neural networks can be used to fill the gap between ESM and observations[26]. Computational infrastructure using ML/AI that links ESM and observational data are needed that improve the state-of-the-art in several ways: a) parameter estimation; b) data assimilation—once the ML/AI models are trained with initial observational data (along with ESM data or within ESM workflow) they can be re-trained on the fly as new observational data comes in[27]. As aspects in a) and b) are learned by the ML/AI model; it can help climate scientists with experimental design; for instance, it can help in identifying where the next sensor should be ideally placed.

Through such methods, the ESM-predicted extreme events (i.e., maximum precipitation) distribution can be extracted, and this distribution can be compared to observations to see that extremes are accurately captured in the ESM[28]. Once the ESM can be downscaled, trained or tuned to the assimilated observations of extremes, and ML/AI- assisted physical extreme events characterization will be vastly improved, and the intensity, frequency, magnitude, and reoccurrence of such events can be looked at within the model directly, rather than through a cascade of downscaled model outputs and secondary land surface and/or physical models. ML/AI approaches embedded within ESMs will allow for researchers to move more quickly to action on both mitigation and adaption measures, to protect society and economy from the deleterious effects and impacts of extreme hydroclimate events.

Characterization of Extreme Hydroclimate Events in Earth System Models using ML/AI

References

1 Field, C. B. *et al. Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. (Cambridge University Press, 2012).

2 Bird, K. J. *et al.* Circum-Arctic resource appraisal: Estimates of undiscovered oil and gas north of the Arctic Circle. Report No. 2327-6932, (US Geological Survey, 2008).

3 Larsen, P. H. *et al.* Estimating future costs for Alaska public infrastructure at risk from climate change. *Global Environ. Change* **18**, 442-457 (2008).

4 Markon, C., S. Gray, M. Berman, L. Eerkes-Medrano, T. Hennessy, H. Huntington, J. Littell, M. McCammon, R. Thoman, and S. Trainor. in *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II* (ed D.R. Reidmiller, C.W. Avery, D.R. Easterling, K.E. Kunkel, K.L.M. Lewis, T.K. Maycock, and B.C. Stewart) 1185–1241 (U.S. Global Change Research Program, 2018).

5 Melvin, A. M. *et al.* Climate change damages to Alaska public infrastructure and the economics of proactive adaptation. *Proc. Natl. Acad. Sci.* **114**, E122-E131, doi:10.1073/pnas.1611056113 (2017).

6 Williams, J. W. & Jackson, S. T. Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment* **5**, 475-482, doi:10.1890/070037 (2007).

7 Bennett, K. E. & Walsh, J. Spatial and temporal changes in indices of extreme precipitation and temperature for Alaska. *International Journal of Climatology* **35**, 1434-1452, doi:10.1002/joc.4067 (2015).

8 Bennett, K. E., Cannon, A. J. & Hinzman, L. Historical trends and extremes in boreal Alaska river basins. *J of Hydro* **527**, 590-607, doi:doi.org/10.1016/j.jhydrol.2015.04.065 (2015).

9 Solander, K. C., Bennett, K. E. & Middleton, R. S. Historical shifts in streamflow extremes in the Colorado River Basin. *Journal of Hydrology: Regional Studies.* **12** (2017).

10 Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society* **93**, 485-498, doi:doi.org/10.1175/BAMS-D-11-00094.1 (2012).

11 Wehner, M. F., Smith, R. L., Bala, G. & Duffy, P. The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. *Climate dynamics* **34**, 241-247 (2010).

12 Schiermeier, Q. The real holes in climate science. *Nature* **463**, 284-287 (2010).

13 Reichstein, M. *et al.* Climate extremes and the carbon cycle. *Nature* **500**, 287-295, doi:10.1038/nature12350 (2013).

14 Zhang, X. & Zwiers, F. W. in *Extremes in a Changing Climate* (eds Amir AghaKouchak *et al.*) 1-14 (Springer, 2013).

15 McGuire, A. *et al.* in *Land Change Science* 139-161 (Springer, 2012).

16 Bennett, K. E., Talsma, C. & Boero, R. Concurrent Changes in Extreme Hydroclimate Events in the Colorado River Basin. *Water Special Issue "Hydrological Extremes in a Warming Climate: Nonstationarity, Uncertainties and Impacts"* (2021 in review).

17 Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A. & Pulwarty, R. S. Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resources Research* **35**, 2145-2160 (1999).

18 Vose, R., Menne, M., Durre, I. & Gleason, B. in *AGU Spring Meeting Abstracts.* A41A-08.

Characterization of Extreme Hydroclimate Events in Earth System Models using ML/AI

19 Mather, J. H. & Voyles, J. W. The ARM Climate Research Facility: A review of structure and capabilities. *Bulletin of the American Meteorological Society* **94**, 377-392 (2013).

20 Vesselinov, V. V., Mudunuru, M. K., Karra, S., O'Malley, D. & Alexandrov, B. S. Unsupervised machine learning based on non-negative tensor factorization for analyzing reactive-mixing. *Journal of Computational Physics* **395**, 85-104 (2019).

21 Alexandrov, B. S. & Vesselinov, V. V. Blind source separation for groundwater pressure analysis based on nonnegative matrix factorization. *Water Resources Research* **50**, 7332-7347 (2014).

22 O'Malley, D. & Vesselinov, V. A combined probabilistic/nonprobabilistic decision analysis for contaminant remediation. *SIAM/ASA Journal on Uncertainty Quantification* **2**, 607-621 (2014).

23 Vesselinov, V. V., Alexandrov, B. S. & O'Malley, D. Contaminant source identification using semi-supervised machine learning. *Journal of contaminant hydrology* **212**, 134-142 (2018).

24 Karra, S., O'Malley, D., Hyman, J. D., Viswanathan, H. S. & Srinivasan, G. Modeling flow and transport in fracture networks using graphs. *Physical Review E* **97**, 033304, doi:10.1103/PhysRevE.97.033304 (2018).

25 Ahmmed, B., Mudunuru, M. K., Karra, S., James, S. & Vesselinov, V. V. A comparative study of machine learning models for predicting the state of reactive mixing. *Journal of Computational Physics*, 110147 (2021).

26 Gu, J. *et al.* Recent advances in convolutional neural networks. *Pattern Recognition* **77**, 354-377 (2018).

27 Wang, S., Karra, S. & O'Malley, D. Large-scale Inversion of Subsurface Flow Using Discrete Adjoint Method. *arXiv preprint arXiv:1906.01132* (2019).

28 Talsma, C., Bennett, K. E. & Vesselinov, V. V. Characterizing Extreme Events using Unsupervised Machine Learning Methods for Improved Understanding of Future Droughts in the Colorado River Basin. *In preparation for special issue of JGR Solid Earth, G-Cubed, Tectonics and Earth and Space Science, "Machine learning for Solid Earth observation, modeling and understanding"*. (2021).

Suggested Partners/Experts

Dr. Dan Cooley, Professor, Colorado State University, https://www.stat.colostate.edu/~cooleyd/

Dr. Alex Cannon, Research Scientist, Environment Canada, https://www.eoas.ubc.ca/people/alexcannon

Dr. Amir AghaKouchak, University of California, Irvine, http://amir.eng.uci.edu/index.php